EXECUTIVE SECRETARIAT

ROUTING SLIP

TO:

		ACTION	INFO	DATE	INITIAL
1	DCI				
2	DDCI				
3	EXDIR				
4	D/ICS				
5	DDI		X		
6	DDA				
7	DDO		X		
8	DDS&T				
9	Chm/NIC				
10	GC				
11	IG				
12	Compt				
13	D/OLL				
14	D/PAO	X			
15	D/PERS				
16	VC/NIC				
17					
18					
19					
20					
21					
22					

SUSPENSE ______________________ Date

Remarks

TO #14: For direct response as appropriate.
(ASG/DI might be interested.)

STAT

Executive Secretary

22 NOV 85

Date

ER

3637 (10-81)

Executive Registry
85- 4625

STAT [redacted]

15. November 1985

Mr. William Casey
Director
Central Intelligence Agency
Washington DC 20505

Dear Mr. Casey,

Does your agency use election data to analyze political developments in foreign countries? I wish to share with you a new method of analyzing election data that can give long-range indications of major political changes: changes that may include the emergence of a new political party or social movement, a civil war, or the transformation of a democratic government into a totalitarian regime. The method can be applied to any country with fair elections, such as Japan, India, Turkey, or Greece, among many others.

I have enclosed a recent journal article that shows how this method works with American election data. I have found similar results in Japan, India, Indonesia, Weimar Germany, and other nations. The method is based on a new scientific theory of voting behavior, which relates voting to the structure of a society and certain universal characteristics of how the human brain processes information. The method may look complex, but it is easy to apply and very inexpensive.

Although the ideas here are new and will undergo further refinements, I believe that the time is right to begin a concerted program that will collect and analyze election data from all nations. Over a period of years, this will prove to be a valuable complement to traditional short-term political intelligence gathering.

Should you or your agency be interested in pursuing these ideas, I will be available to assist. Enclosed is my resume and that of my associate [redacted] We also have considerable experience in evaluating the usefulness of information to decision makers, which would be an important adjunct to a program such as this.

STAT

Sincerely,

Stephen Coleman

Stephen Coleman, Ph D

DCI EXEC REG

P-364

STAT

Page Denied

Next 5 Page(s) In Document Denied

Approved For Release 2009/09/21 : CIA-RDP87M00539R002804720026-8

VOLUME 11/NUMBER 2/1985
ISSN 0022-250X
JMTSBT 11(2) 95–189 (1985)

The Journal of MATHEMATICAL SOCIOLOGY

JOURNAL OF MATHEMATICAL SOCIOLOGY

CONTENTS Volume 11, Number 2

GORDON AND BREACH SCIENCE PUBLISHERS New York • London • Paris • Montreux • Toky

Approved For Release 2009/09/21 : CIA-RDP87M00539R002804720026-8

Journal of Mathematical Sociology, 1985, Vol. 11, pp. 95-130
0022-250X/85/1102-0095$30.00/0

Printed in the United States of America

THE HUMAN BRAIN, SOCIAL CONFORMITY, AND PRESIDENTIAL ELECTIONS

STEPHEN COLEMAN
Minneapolis, Minnesota

This paper investigates aggregate voting behavior in American presidential elections from 1904 to 1980; the state is unit of analysis. We abandon the traditional assumption that voting is a rational process and instead construct a model to test the power of social conformity over voting decisions. The model, which uses the entropy measure of statistical information theory, allows specific numerical tests of its validity and, additionally, reveals unexpected dynamic patterns in voting behavior. We find four distinct processes related to social conformity that affect voter turnout and the distribution of votes among presidential candidates. A mathematical model for each process is identified and parameters are estimated from election data.

FOREWORD

Rationality or conformity—which better describes people's voting behavior? Here we cast our vote for conformity after an analysis of American presidential elections from 1904 to 1980.

Over the past few decades we have witnessed an extensive development of voting theory that has at its root the assumption that voting is an individualistic, rational act. Indeed much of voting theory has become the province of economists, who see voting choices as analogous to economic choices, where rationality is axiomatic. One certainly cannot refute the possibility that individual voters evaluate and rank political candidates according to some logical preference scheme. This possibility is the basis of our faith in the democratic electoral process. But is our faith well founded in empirical evidence?

Here we shall take a fresh look at voting, with new assumptions. These will lead us to mathematical models of voting that hold up under empirical testing and reveal unexpected patterns in collective voting behavior—patterns which cannot be explained by the conventional thinking about voter rationality. We shall demonstrate that voting has to do with conformity rather than with rationality and that the outcome on an election has more to do with how the human brain starts at a very young age with little further age development, nor is it much related to a person's intelligence, education, culture, or socioeconomic status.

This line of research makes it seem very likely that people have in their minds automatically and at all times information on the relative frequencies of well-known political categories such as defined by political party identification or voting participation.

Assume now that we have a society where people are good estimators of relative frequency. Can they perceive how conformist their society is? Not directly, I maintain. Suppose we ask one of the citizens which of two situations would feel more conformist: first, the case where 70 percent of the people are in one group and 30 percent in another; or case two, where 80 percent are in one group, 15 percent are in a second group, and 5 percent are in a third group. To answer this, I believe, requires more than the knowledge of the probabilities—it requires a calculation or assessment of the balance of probabilities among the various groups. How might the human mind rank the degree of conformity where probabilities and numbers of groups differ?

Again, psychology offers insights. Research has shown that there is a limit on how many groups, or choices, and their relative frequencies that the brain can handle at one time (G. Miller, 1956; H. Miller and Bieri, 1963). The measure for this degree of complexity is the entropy concept of statistical communications theory (Shannon and Weaver, 1969; Fano, 1961). Apparently, this measure captures something important about how the brain processes probabilistic information. Entropy—which we define below—is the inverse of conformity. In other words, there is a limit on how much nonconformity people can perceive.

This limit is a number on a ratio scale and corresponds to the situation where a person has to distinguish accurately among seven or eight equally likely possibilities. (Different experiments give slightly different results.)

The origin of the entropy measure is in communications theory, where it describes the capacity of a communications channel and the structural content of messages being sent and received. So our discussion of conformity and its perception leads us to the view that, in some instances, behavior is a form of communication as the brain perceives it. And because entropy is a unique measure for the capacity of a communication channel (such as the brain), we ought to consider how entropy might relate to social behavior, generally, at levels of structure below the brain's capacity limit.

Before going further, we must clarify a point that might be confusing to social scientists. A usual research practice is to start with a concept, such as "conformity," and then "operationalize" it, that is, find a quantitative measure that seems to encompass what we mean by the concept. Then one tests out hypotheses involving the concept by its surrogate measure. The measure may, of course, have no necessary or unique relation to the concept.

Information theorists (Fano, 1961; Khinchin, 1957; Shannon and Weaver, 1969) argue that in order for a quantity to be a reasonable, consistent, and useful measure of information or uncertainty for a set of alternative events, the measure must have certain axiomatic properties. They then prove that the exact form of the equation above is necessary and sufficient for a measure that has the desired properties.

Entropy is synonymous with "average information" (Fano, 1961), because it is an average of terms of the form $-\log p$, each of which measures the information in the occurrence of an event expected with probability p. Entropy can also be interpreted as a measure of uncertainty (Khinchin, 1957, p. 7). Entropy describes the amount of information that has to be known to remove all of the uncertainty about what will happen in a given situation. The terms entropy, average information, and uncertainty will be used interchangeably. Note, however, that entropy in communications theory is not the same as the entropy of thermodynamics, although both have the same mathematical form. In thermodynamics entropy is a measure of disorder. Attempts have been made to apply this concept of entropy to social analysis as well.

Entropy has the property (by definition) that it is greatest when the probabilities of all n events are equal (to $1/n$). In this case, $H(S) = \log n$. Observe also that as the probabilities become closer to equality, entropy increases up to the maximum. Entropy is never less than zero.

ELECTION ENTROPY HYPOTHESIS

Let us now turn our thoughts to voting. In any election there are two choices that face the voter: whether or not to vote, and, if voting, for whom. One can measure the entropy or uncertainty in each of these two choice situations by computing first the probabilities of the alternatives and then the entropy for each set of alternatives. We may not know the probabilities prior to the election, but we can use the election outcome to calculate retrospectively the probability of a randomly chosen person voting or abstaining or, if voting, the probabilities of voting for the various parties or candidates (for a single office).

We can interpret an election entropy measurement as the degree of difficulty or sense of uncertainty that an arbitrarily chosen citizen—or political observer—in a society will have in trying to predict who will vote for whom. In a high entropy society it will be relatively difficult to predict how others will vote; in a low entropy society we will be able to anticipate more often the random voter's behavior. Thinking of conformity, we see that in a high entropy society people will be relatively less conforming—there will be more factions or larger minorities—as compared to a low entropy society.

People in the society ought to be able to detect changes in the level of conformity through their perceptions of probabilities or relative frequencies of how others will vote, just as we scientific observers can detect change through election

We have no data on how persons behave given that they perceive a certain level of uncertainty or conformity in the voting behavior of their fellow citizens. Yet because a voter can choose only one alternative from a set, we will never be able to assess the degree of conformity or influence of entropy in a society by looking at the behavior of any single individual. We have to look at how individual behaviors combine. The best and only way to examine the influence of entropy on collective voting patterns is to present results that strongly imply the existence of a link between H(P) and H(T) or which show that these measures separately have a strong impact on election outcomes.

We come then to the hypothesis that we want to test first: that the two measures H(T) and H(P) are equal if we take into account the number of choices in each set of alternatives. To put both measures on the same (ratio) scale, one need only divide each by the maximum possible for the set. For the two-choice situation of voting or abstaining, we divide H(T) by log 2; for the n-choice situation of voting for one of the political parties or candidates, we divide H(P) by log n. Thus, the hypothesis is that

$$H(P)/\log n = H(T)/\log 2.$$

As log 2 = 1

$$\boxed{H(P) = \log n\ H(T)} \quad (1)$$

We shall call this predicted relation (Eq. 1) the "entropy (or uncertainty) hypothesis." If true, it implies that people acting together adjust their voting behavior so as to be consistent in how they respond to perceiving the same (relative) level of entropy. In terms of conformity, the hypothesis is that people are making their voting decisions as if society exerts the same degree of conformity in each of the two behavioral situations. If a society is conformist to a certain degree in getting people to vote, it will be conformist to the same degree in how the vote divides among the political parties.

Several additional points bear on the testing of the hypothesis. First, the number of political parties must be determined. It may not be obvious how many choices people generally perceive in an election. So this becomes an empirical question that we must investigate.

A second point is that one would, it seems, want to examine those elections where the population was most aware that an election was taking place and knowledgeable about the different political parties, that is, able to distinguish them at least. This consideration leads us to examine the major national elections, such as American presidential elections or parliamentary elections in most other countries.

A third difficulty with the hypothesis is that it says nothing about change over time. The presence of an integer n in Eq. (1) indicates that we have to deal with discrete change, as when the number of parties changes, in addition

A second test is that the shape of the distribution of states arranged by H(P) and turnout ought to be nearly parabolic and, specifically, be symmetrical about the line where turnout equals 50 percent. In this instance, we use parabolic (second degree polynomial) curve fitting and test whether in fact a parabola is appropriate, and if so, whether it has a maximum and, if a maximum, whether it is near 50 percent turnout.

A third empirical test is whether the H(P) values exceed the psychological limit on entropy capacity at about 3 bits. The United States does not present a good case for examining this, but we shall refer to prior results on this point in two other countries.

To my knowledge this is the only social theory that makes specific numerical predictions of such generality. We shall examine additional numerical and statistical tests below when we turn to the dynamic properties of entropy change.

ENTROPY VERSUS RATIONALITY

The purpose of an election, at least in the democratic view, is to give voters a choice. Presumably each voter will consider the merits of candidates or parties and rank them, picking out the one best suited to his interests. The rational model of voting starts with this assumption. Economists have been strongly involved in this branch of voting theory because of its analogy to the economic decision making of the consumer in the marketplace (Arrow, 1963; Downs, 1957).

The assumption of rationality is also at the heart of most voting research by political scientists, although the assumption may not always be explicitly stated. Authors who do make a clear exposition of the implications of rational voting include, among many others, Campbell et al. (1966), Davis et al. (1970) and Kelley (1983).

The approach taken by proponents of rational voting theory makes sense because the premise of rationality is central to our ideas about democratic government. But is our faith well-founded? Do the mathematical models of rational voting theory have anything to say about what actually happens in an election?

Before presenting the results of testing the entropy hypothesis on American presidential elections, let us consider what it means for the interpretation of the voting act. That the distribution of votes among parties or presidential candidates is already determined to a great degree by the voter turnout implies that nonvoters substantially determine the results of the election. This is outrageously opposed to any interpretation of rationality in elections or the preference ranking of candidates by the voters. Therefore, I maintain, the entropy hypothesis is a good test of the rational model of voting itself.

We can also think of H(P) as a measure of the competitiveness of the election. If political parties garner more equal shares of the vote, H(P) increases. If there is greater except for states in the American South. Only the South shows the expected pattern of turnout increasing with competitiveness. This result partially supports the entropy hypothesis because turnout in the South has been consistently below 50 percent. That she does not find a strong relationship in other states may have to do with the level of turnout (which is not reported) or with the choice of elections. We will show that strong time-lagged effects exist which have not been recognized in any of the prior election analyses, and this neglect has worked against detection of the turnout-competitiveness link in past research.

Ashenfelter and Kelly (1975) take a broader look at the question of voter turnout in presidential elections and attempt to weigh rational or "economic" voting assumptions, such as costs and benefits perceived by the voter, against other interpretations of why people vote. Their analysis combines survey data for presidential elections in 1960 and 1972 with data on certain system-level characteristics. The topical variety of factors included in their analysis is shown by this partial list: personal characteristics of voters, costs of voting, strategic value of voting, partisanship of voters, past party competitiveness, interest in the campaign, and sense of obligation to vote.

They explicitly reject the idea that voters weigh the costs and benefits of voting and conclude that it is the "sense of duty or obligation as the primary motivation for voting" (p. 724). And they further state that (p. 724), "Our results...lead us to conclude that most voters correctly perceive that their vote will have no effect on the outcome of a presidential election..." They do not find a strong or consistent link between party competitiveness and turnout.

I believe that their method of analysis and choice of elections (as we see below) mitigated against finding a turnout-competitiveness relationship. Nevertheless, their findings seem to lend more credence to a conformity-based theory of voting participation than to the rational theory.

PAST RESEARCH RESULTS

In my earlier book (Coleman, 1975), I examined the entropy hypothesis for the United States, India, Japan, and Germany—all with subnational units of analysis. Elections were selected over a wide range of time: United States presidential elections in 1920, 1924, 1940, 1948, and 1968; India, the parliamentary election of 1962; Japan, thirteen parliamentary elections from 1924 to 1937 and 1949 to 1967; and Germany, parliamentary elections of 1924 to 1933. The entropy hypothesis was generally well supported in this extensive analysis; but in countries where there are different numbers of parties in different constituencies or where there are rapid changes over time, it was necessary to modify the hypothesis. That is, it became apparent that one had to know something of the dynamics of $H(P)$ and $H(T)$ change in order to explain fully the observed patterns. So it is that we undertake a study of entropy dynamics here.

Inspection of maps of H(P) values in the United States in 1968, or other recent elections, also indicated the presence of a diffusion of voting conformity that tended to smooth out differences between neighboring states, although not for the country as a whole.

Spatial H(P) distributions in the United States have the look of daily weather maps that plot lines of equal temperature across the country. Such a pattern is characteristic of a certain class of mathematical functions known as harmonic or potential functions (Garabedian, 1964; Kellog, 1953). This class of functions appears in diverse physical sciences including diffusion processes. Harmonic functions have unique properties that can be tested for empirically, which we shall do below.

We also had observed in the elections of Weimar Germany certain time lags in the change of several entropy-related variables with respect to one another. This raised the possibility that large changes in H(P)—as happened in the fall of the Weimar Republic and the rise of Hitler's regime—might be predicted years in advance by detecting changes in antecedent variables. As we now turn to the testing of the entropy hypothesis, evidence for the existence and potential use of time lag data becomes compelling.

TESTING THE ENTROPY HYPOTHESIS

Our data set comprises election data for American presidential elections—the vote for president—from 1904 to 1980 for each state. (We excluded Alaska, Hawaii, and the District of Columbia a priori so that the same states might be followed over time; but in 1904, there were only 45 states and in 1908 only 47.) Data were extracted from readily available references, including *Statistical Abstract of the United States* and *Historical Statistics of the United States, Colonial Times to 1970*. For each state for each election, we have three values: H(P), H(T), and turnout.

Because data on American elections are so readily available and because, for the most part, our calculations are so easily done, it will be possible for many readers to duplicate these results without difficulty.

The best test of the entropy hypothesis would be a presidential election where all potential voters would agree on the number of candidates competing for the presidency. If a party obtains only, say, 1 percent of the vote, it is likely that many voters are unaware of its existence and so cannot estimate the relative frequencies of people voting for its candidate; it will not be a part of their entropy perception. When the size of a party is larger, at 5 percent or 10 percent of the vote, it is likely that many but not all potential voters will have some experience by which to estimate its share of the vote. In such an election, we must anticipate that the testing of the hypothesis will give us an intermediate coefficient value, between values of log n.

Elections such as 1912, 1924, or 1968 when minority parties were widely recognized throughout the country ought to

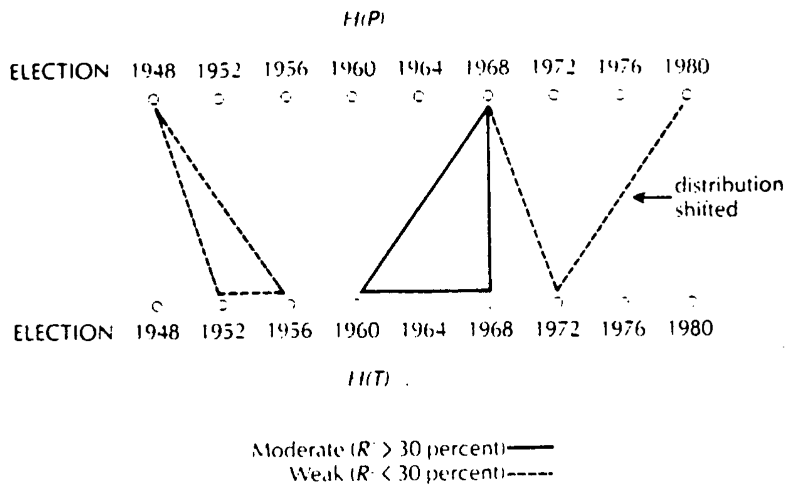

Figure 4. Time lag relationships between H(P) and H(T) for the model H(P) = cH(T) from 1948 to 1980.

It may be that we will have to wait for 1984 or beyond to see the link between H(T) and H(P) completely reestablished.

The strongest connections between H(P) and H(T)—and these are all positive relations—as measured by R^2 are: 1912 H(P) and 1916 H(P) with 1920 H(T), R^2 = 70 percent; 1920 H(T) and 1932 H(P), R^2 = 71 percent; 1920 H(T) with 1940 H(P), R^2 = 81 percent; and 1924 H(T) with 1944 H(P), R^2 = 79 percent.

The pattern of time delays is also borne out if we examine *changes* (differences) from one election to the next in H(P) and H(T). We observe that change in state H(P) levels from 1908 to 1912 accounts for 43 percent of variation in H(T) change from 1916 to 1920; change in H(T) from 1920 to 1924 in turn accounts for 49 percent of variation in H(P) change from 1928 to 1932. This last difference in H(P) then explains 32 percent of variation in H(T) change from 1940 to 1944. Later in the series, we find that change in H(T) from 1956 to 1960 accounts for 59 percent of the variation in H(P) changes between 1964 and 1968.

In short, we see that changes in H(P) or H(T) have a reciprocal effect on one another with a time delay of about 2 to 3 elections or 8 to 12 years.

For the other pairs of elections, the strength of the H(T)-H(P) relationship drops off, although in many intervals only gradually. For example, H(T) from 1960 to 1968 show almost the same, if slightly decreasing, strength of relationship to 1968 H(P). The explanation of what is happening during the periods of weaker H(P)-H(T) linkage must await our analysis of their respective dynamics.

We next consider how well the H(P)-H(T) relationship predicts the expected value log n for n political parties. That is, we estimate the coefficient c in the regression of H(P) on H(T).

the maximum is at 66 percent (R^2 = 78 percent). For turnout at 1920, by contrast, the maxima align themselves much more closely to 50 percent, regardless of the choice of year for H(P). This change in symmetry alignment at 1920 may simply reflect time delays, or it may be the effect of the introduction of women's suffrage in presidential elections beginning at 1920. The "errors" in the parabolic alignment before 1920 may in fact show that women did affect voting behavior even though they were not counted as potential voters in the turnout computation. (A similar effect of shifted symmetry was observed in our previous analysis of Japanese prewar elections when suffrage gradually expanded among males (Coleman, 1975).)

The parabolic regressions that have the best fittings of H(P) to turnout are of nearly the same strength and location as the strongest H(P)-H(T) relationships cited above. We give these examples: 1932 H(P) fit to 1920 turnout has a maximum at 56.9 percent turnout with R^2 = 84 percent; 1940 H(P) fit to 1920 turnout has a maximum at 57.0 percent turnout with R^2 = 86 percent; 1932 H(P) fit to 1924 turnout has its maximum at 56.9 percent turnout with R^2 = 79 percent; and 1940 H(P) fit to turnout in 1924 has a maximum at 54.6 percent turnout with R^2 = 80 percent.

Other examples of parabolic fitting are: 1920 turnout with 1920 H(P), maximum at 50.6 percent turnout and R^2 = 36 percent; 1924 turnout with 1920 H(P), maximum at 48.8 percent turnout and R^2 = 32 percent; 1924 H(P) and 1924 turnout, maximum at 59.2 percent with R^2 = 55 percent; 1928 H(P) and 1928 turnout, maximum at 52.1 percent turnout and R^2 = 35 percent.

In recent elections, the low turnout rates in the South have increased to a level (nearer 50 percent) where the parabolic shape is not much in evidence. In 1968, for instance, H(P) values decrease linearly with increasing turnout, which is not less than 44.0 percent in any state. (R^2 = 44 percent for a linear regression.)

One can see from these results that the entropy hypothesis is well-established, but that voting behavior related to the H(P)-H(T) link is not always the dominant process affecting the election outcome. Indeed we shall see that the H(P)-H(T) process is but one of four (or perhaps more) autonomous dynamic processes involving entropy-related voting behavior.

The time-lagged effects are as interesting as they are unexpected. Clearly, they too weigh against the rationalist idea that voters are examining and ranking the parties, candidates, and issues of the moment. Beyond that, one must suppose that the time delays represent the length of time it takes for changes in H(P) or H(T) to propagate through the population and for people to adjust the conformity in one voting situation to that of the other prior situation. Obviously, it takes some years for a new political party to emerge in a country as large and diverse as the United States, given that an increase in H(P) is anticipated.

Time delays also suggest that a society "remembers" its entropy level. The social structure—the relationships among people and their expected behavior—captured by an entropy of all the processes into a single model. The models that we introduce here, however, also lend themselves to a computer simulation, which can give insights into the combined effect of the separate dynamics on system entropy levels. A simulation approach permits consideration of system stability questions as well.

THE H(P) DYNAMIC

If one scans average H(P) values for the states over the period from 1904 to 1980 (Table 1), it is evident that

TABLE 1. Average H(P) values for the States - 1904 to 1980.

Year	H(P) bits
1904	1.12
1908	1.16
1912	1.62
1916	1.12
1920	1.08
1924	1.26
1928	0.96
1932	0.98
1936	0.97
1940	0.93
1944	0.95
1948	1.14
1952	0.97
1956	0.97
1960	0.99
1964	0.94
1968	1.36
1972	1.04
1976	1.15
1980	1.25

the H(P) level goes up and down over time. A closer look shows also that after an election with a large H(P) increase will invariably come an election with a large H(P) decrease. After 1912, a high point, we find a relatively low level of H(P) in 1916—the lowest over several elections up to that time. After 1924 follows another low point in 1928. The same is true of 1948 to 1952 and 1968 to 1972. In each of these four cases, an election with a strong third-party movement was followed by an election with little or no third-party activity. Did the third party "burn out"? Did its supporters give up in the next election after suffering defeat? What explains this?

The ebb and flow of party strength has been noted by a number of voting analysts. Stokes (Campbell et al., 1966, Chap. 10) investigated whether the fluctuations in party dominance in American presidential elections fit a random (walk) change in velocity is the acceleration from time t to t + 2. The model of oscillatory behavior is a linear relation between acceleration and level of H(P)

$$\Delta_t^2 H(P) = c_1 H(P)_{t+1} + c_2 \quad (2)$$

where c_1 and c_2 are arbitrary constants. This difference equation will have oscillatory behavior, that is, the solution will be oscillatory, for c_1 in the interval (-4,0). (Goldberg, 1958; the auxiliary equation, p. 138, has complex roots in this interval.)

If $c_2 = 0$, then the oscillation centers about the point $H(P) = 0$. If $c_2 \neq 0$, then the motion is about $H(P) = c_2/c_1$, which is where $\Delta^2 H(P) = 0$.

A regression analysis of state $\Delta_t^2 H(P)$ levels against H(P) levels for each election t will give us estimates of c_1 and c_2. We have in fact done this analysis for each election from 1908 to 1976.

This regression analysis, which tests Eq. (2), has a weakness in that the strength of the relationship is hard to interpret. This is because H(P) appears (implicitly) on both sides of the equality. To remedy this problem, and also give another perspective on the oscillatory process, we perform a second regression of $\Delta_t H(P)$ on $\Delta_{t-1} H(P)$ values for the states at each election; that is, we carry out cross-sectional analysis of change immediately before and after each election. The regression model is

$$\Delta_t H(P) = c_1 \Delta_{t-1} H(P) + c_2 \ . \quad (3)$$

This equation has oscillatory behavior for c_1 negative. (Goldberg, pp. 136-7; the auxiliary equation has only real roots, which are c_1 and 1.) By this method, and by looking at plots of $\Delta_t H(P)$ against $\Delta_{t-1} H(P)$, one can show directly the up and down—down and up—motion of the states over time.

The first of the two regression analyses (Eq. 2) shows that the oscillatory model holds most strongly at t = 1912 (R^2 = 73 percent), 1924 (R^2 = 69 percent), 1932 (R^2 = 71 percent), 1936 (R^2 = 29 percent), 1948 (R^2 = 83 percent), 1956 (R^2 = 78 percent), 1968 (R^2 = 76 percent), 1972 (R^2 = 79 percent), and 1976 (R^2 = 34 percent). (In 1956, Mississippi was excluded as an outlier.)

In each of these elections, the constant c_1 has a negative value, as expected, and c_1 ranges from -.43 in 1936 to -1.95 in 1968, for an average of -1.25 over the nine elections.

The center of the oscillation, estimated from the values of c_1 and c_2, ranges from 0.65 in 1912 to 1.27 in 1968 and averages 0.95 over the nine elections.

The second of the regression analyses, testing the relationship between first differences (Eq. 3), gives results similar to the above but with lower R^2 values, as we might expect. The strongest relation between differences occurs when the middle election of the three is at t = 1912 (R^2 = 88 percent), 1920 (R^2 = 43 percent), 1924 (R^2 = 46 percent), 1928 (R^2 = 32 percent), 1936 (R^2 = 28 percent), 1948 (R^2 =

of alternatives is possible, we do not observe oscillatory behavior.

Yet oscillatory behavior is also related to the concept of social "memory." If people acting collectively try to return their behavior to a previous level, then they must have a memory of it. Or they may have a memory of the rate of change.

THE H(T) DYNAMIC

In physical systems, we find many examples of systems with a "memory." These are oscillatory systems where the memory is in the form of stored or potential energy.

There are also many physical processes or systems that seem to act without memory. Examples of these are radioactive decay or, at times, population growth. Such processes are modeled as exponential functions of time. What happens in these processes from one moment to the next depends only on the level of the system at the instant, not on what happened in the past.

The behavior of H(T) in American presidential elections shows signs of being a memoryless process. Here we tested the model

$$\Delta H(T)_t = c_1 H(T)_t + c_2 \tag{4}$$

where c_1 and c_2 are arbitrary constants. The sign of c_1 is critical in this type of system: a positive sign causes exponential growth; a negative sign with c_1 in the interval (-1,0) gives monotone convergence (Goldberg, pp. 77-79) to the point $H(T) = -c_1/c_2$ where $\Delta H(T) = 0$ and the system stops changing. The greater the initial difference between where the system is now and its final finite destination, the faster it will be changing at that moment.

We must be careful in this analysis to rule out the trivial or spurious. This is a danger here because of the limit of H(T) values at 1 bit. For example, if many states increased H(T) to the limit between elections, those states would automatically satisfy the model (Eq. 4). As direct inspection of the data on $\Delta H(T)$ reveals, however, it is unusual for states, except perhaps if they are very near 1.0 already, to increase to the limit. (This corresponds to a change in turnout to 50 percent.)

Applying the exponential model, which was selected after an investigation of many alternatives, we find that it holds pretty well for positive $\Delta H(T)$ in elections prior to 1960. After 1960, it holds for negative $\Delta H(T)$ as well, with the result that the relationship is captured best by the absolute value of $\Delta H(T)$ as in

$$|\Delta H(T)_t| = c_1 H(T)_t + c_2. \tag{5}$$

In sum, the exponential model seems to have become stronger or, rather, this type of entropy-related behavior seems to have become more prevalent, in recent elections as compared to earlier elections.

It may also be that people can use the equal probability situation (1 bit) for two choices as a fixed reference point in comparing or adjusting what they perceive as the level of social uncertainty.

SPATIAL ANALYSIS

A final piece to the voting behavior puzzle is the spatial dimension. The spatial properties or dynamics are the result of the diffusion of voting behavior patterns among neighboring states. That such a diffusion takes place is strongly shown by certain large-scale spatial patterns in voting entropy distributions.

Our earlier analysis of the United States had shown that the spatial distribution of state H(P) values has the characteristics of harmonic or potential functions, which in physical systems may result from diffusion or propagation effects. Harmonic functions have unique properties that one can test for on spatial entropy data (Garabedian, 1964; Kellog, 1953).

Two-dimensional spatial analysis is very difficult, however, and so we have adopted a method to reduce the analysis to one of a single spatial dimension. We shall "linearize" the geography of American states.

To convert a two-dimensional map of the United States to a one-dimensional "map," we order the states such that each state in the list of states shares a geographic border with the state immediately above and below it on the list. The procedure begins with Washington, Oregon, California, and continues on to Maine. The entire arrangement of the states by this ranking method is in the Appendix.

The idea of the linearization is that the diffusion of voting behavior, if it exists, ought to be stronger over close distances—as from one state to a neighboring state—than over greater distances. Thus, the H(P) values of states in the list ought to be more influenced by H(P) values in adjacent states than by H(P) in more remote states, which will generally be farther away in the list. This method will, of course, lose some information on spatial relationships that a two-dimensional analysis might preserve. Nonetheless, we will see that this method works quite well. (The method might be applicable to other spatial problems in social analysis; but to my knowledge, this is the first try at it.)

A harmonic function is one that satisfies the Laplace equation, which in the one-dimensional case for H(P) is

$$\Delta_x^2 H(P) = 0 \tag{7}$$

for (second) differences in the spatial dimension x. If the nth state in the list of states has value $H(P)_n$, then the Laplace equation becomes

$$\Delta^2 H(P)_n = 0$$

The spatial patterns also show the importance of conformity in election outcome, especially at a close geographic range, and that it is conformity on an entropy dimension that concerns people in their voting choices.

In comparison with the long time delays observed in the H(P)-H(T) linkage, or the relatively long period of the oscillatory motion of H(P), the spatial process must go on more rapidly. This one can infer from how quickly the fit of the spatial regression model (Eq. 8) can change from one election to the next, especially for H(P). (Indeed, analysis of over time data shows only a weak or intermittent presence of spatial factors between elections.)

The spatial effect of H(P) is also additive with respect to the H(T)-H(P) linkage. If one includes both models in the same regression, as for 1968, the total explained variance increases to about 80 percent.

Other properties of a two-dimensional harmonic function can also be used for tests that such a spatial pattern exists. These properties include the fact that maximum and minimum values for a two-dimensional harmonic function must fall along the boundary of the system—as in a border or coastal state—and that averages around concentric circles must be equal. A test for this latter property might be for equality of average H(P) or H(T) values in interior and boundary states. I leave to the reader further investigations on these lines.

As a practical matter, a strong spatial harmonic pattern is less interesting than the case of a spatial distribution that is distinctly *not* harmonic. A substantial departure from the harmonic distribution implies a lack of national integration and the potential for dramatic change or conflict. In our earlier research, we found two such cases: the United States just prior to the Civil War, and Russia of 1917.

SUMMARY

In Table 2 we present a comprehensive view of which dynamic models apply best to which elections. The criteria for inclusion in Table 2 are based on levels of explained variance (R^2) of the previously discussed regression analyses.

For the H(P) = cH(T) model, Table 2 reports the elections where R^2 exceeds 30 percent and 60 percent for either H(T) or H(P) in a time-lagged relationship. For the oscillatory behavior of H(P), we include the model in Table 2 if either Eq. (2) (the acceleration model) holds with $R^2 > 60$ percent, or if Eq. (3) (the velocity model) holds with $R^2 > 30$ percent. (We recall that the explained variance for Eq. (2) models will tend to run high; so the inclusion level is set more conservatively.)

We also take a conservative approach to including H(T) change. Only cases where the absolute value model (Eq. 5) holds well are included rather than the similar, but less general, results of increasing H(T) only.

For spatial models (Eq. 9), we lower the inclusion level to $R^2 > 20$ percent in order to capture this somewhat weaker dynamic effect. Cases where $R^2 > 60$ percent are also indicated.

levels, a small change in entropy can yield a substantial change in the balance between parties, especially in a two-party election.

But our method strips away the political veneer of elections. As the candidates and issues of one presidential campaign give way to new passions and personalities at successive elections, we observe that millions of people—right through new generations of voters—persist in collective activity that is quite independent of the political transients.

This new interpretation of an election as an entropy measurement brings us tantalizingly close to the joining of cognitive processes—the estimation of relative frequency and entropy in the human mind—with emergent large-scale social and political processes. This is a truly exciting prospect.

Psychological research on frequency estimation shows that elections, including voter participation and party structure, ultimately begin in the automatic processes of the brain rather than in the deliberative, rational side of thinking. Furthermore, entropy is uniquely necessary to the measurement and explanation of voting behavior because it captures how the brain works in assessing levels of conformity for alternative categories of varying likelihood. We have not, of course, discovered all of the steps between perception and action on the individual level, but we see the results in the collective behavior patterns—where indeed it may only be seen.

The findings here cause our interest in elections to shift from traditional concerns such as voter attitudes and campaign issues to broader questions of how a society works. Suddenly, we need to know much more about how a society "remembers," about the origins of social categories, about communication among individuals, and about the accuracy of our perceptions. We also need to interpret the effect of entropy dynamics on the political future.

The past hangs over future presidential elections and national destiny to a much greater degree than one might have imagined possible. Yet, as stifling as this may seem, it holds the promise for looking ahead and for anticipating political disasters such as befell Weimar Germany. By freeing our political destiny from the grip of public opinion, we begin to see that many political catastrophes may not be what people want but are a result of dynamic processes in society that speed on to unintended ends. An economic depression is not the result of a people's desire for unemployment and poverty; neither are war, revolution, repression, and authoritarian regimes inevitably the fruit of popular intent.

"Then what shall we do?"

REFERENCES

Alba, Joseph W. et al. (1980) Automatic encoding of category size information. *Journal of Experimental Psychology: Human Learning and Memory*, 6: 370-378.

Miller, George A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63: 81-97.

Shannon, C. E. and Weaver, Warren (1969) *The Mathematical Theory of Communication*. Champain-Urbana, IL: University of Illinois Press.

APPENDIX I: THE LINEARIZED RANKING OF STATES

1. Washington
2. Oregon
3. California
4. Nevada
5. Idaho
6. Utah
7. Arizona
8. New Mexico
9. Colorado
10. Wyoming
11. Montana
12. North Dakota
13. South Dakota
14. Nebraska
15. Kansas
16. Oklahoma
17. Texas
18. Louisiana
19. Arkansas
20. Missouri
21. Iowa
22. Minnesota
23. Wisconsin
24. Illinois
25. Indiana
26. Michigan
27. Ohio
28. Kentucky
29. Tennessee
30. Mississippi
31. Alabama
32. Florida
33. Georgia
34. South Carolina
35. North Carolina
36. Virginia
37. West Virginia
38. Maryland
39. Delaware
40. New Jersey
41. Pennsylvania
42. New York
43. Connecticut
44. Rhode Island
45. Massachusetts
46. Vermont
47. New Hampshire
48. Maine

APPENDIX II: REGRESSION EQUATION ESTIMATES FOR THE MATHEMATICAL MODELS

MODEL: $H(P) = cH(T)$

Year H(P)	Year H(T)	c (error)	R %	F*	Significance*
1904	1920	1.24 (.03)	34	30	.001
1908	1920	1.28 (.03)	36	36	.001
1912	1912	1.85 (.05)	25	16	.001
1912	1920	1.80 (.03)	70	107	.001
1912	1924	1.84 (.03)	69	102	.001
1916	1920	1.24 (.02)	70	107	.001
1916	1924	1.27 (.02)	62	83	.001
1916	1928	1.30 (.02)	65	95	.001
1920	1920	1.19 (.03)	26	27	.001
1924	1920	1.39 (.03)	52	53	.001
1924	1924	1.41 (.03)	36	37	.001
1924	1928	1.45 (.03)	48	43	.001
1932	1920	1.11 (.02)	71	171	.001
1932	1924	1.13 (.03)	72	143	.001
1932	1932	1.17 (.04)	37	31	.001

Year H(P)	Year T	c_1 (error)	c_2 (error)	c_3 (error)	R^2 %	F	Sig.	Max* at T =
1912	1912	0.051 (.007)	-.00038 (.00007)	0.11 (.15)	76	73	.001	66
1912	1920	0.048 (.006)	-.00041 (.00008)	0.36 (.12)	76	72	.001	59
1916	1916	0.042 (.006)	-.00033 (.00005)	-.079 (.13)	70	53	.001	63
1916	1920	0.036 (.005)	-.00034 (.00006)	0.28 (.09)	69	50	.001	53
1920	1920	0.024 (.006)	-.00024 (.00007)	0.54 (.11)	35	12	.001	51
1920	1924	0.023 (.006)	-.00023 (.00008)	0.61 (.11)	32	10	.001	49
1924	1920	0.032 (.006)	-.00028 (.00007)	0.47 (.12)	55	27	.001	56
1924	1924	0.025 (.007)	-.00022 (.00008)	0.63 (.12)	55	18	.001	59
1928	1928	0.016 (.004)	-.00014 (.00004)	0.61 (.17)	34	11	.001	52
1932	1920	0.043 (.004)	-.00038 (.00005)	-.10 (.08)	84	120	.001	57
1932	1924	0.038 (.005)	-.00033 (.00006)	0.049 (.086)	79	84	.001	57
1932	1932	0.039 (.006)	-.00031 (.00006)	-.14 (.11)	79	84	.001	62
1936	1920	0.039 (.005)	-.00033 (.00006)	-.031 (.089)	79	84	.001	58
1936	1936	0.031 (.006)	-.00022 (.00006)	-.010 (.12)	76	70	.001	72
1940	1920	0.0300 (.003)	-.00027 (.00003)	.17 (.05)	86	134	.001	57
1940	1924	0.028 (.003)	-.00026 (.00004)	0.25 (.06)	80	91	.001	55
1940	1940	0.026 (.003)	-.00020 (.00003)	0.14 (.07)	82	100	.001	67
1944	1924	0.019 (.003)	-.00018 (.00004)	0.52 (.05)	63	38	.001	53
1944	1944	0.021 (.004)	-.00018 (.00004)	0.41 (.07)	63	38	.001	57
1944	1948	0.022 (.004)	-.00020 (.00005)	0.39 (.08)	59	33	.001	56
1948	1948	0.016 (.008)	-.00019 (.00008)	0.89 (.15)	11	2.69	.07	41
1948	1952	0.033 (.010)	-.00031 (.00009)	0.39 (.24)	23	6.6	.003	53

*When the sign of c_2 is negative the model is a parabolic function of T with a maximum at $T = -c_1/2c_2$. The predicted value is T = 50.

MODEL: $|\Delta H(T)_t| = c_1 H(T)_t + c_2$

Year t	c_1	(error)	c_2	(error)	R^2 %	F	Sig.
1904	-0.05	(0.096)	0.060	(0.080)	0	.003	.96
1908	-0.46	(0.06)	0.45	(0.05)	57	56.6	.001
1912	-0.16	(0.07)	0.19	(0.06)	12	5.6	.02
1916	-0.62	(0.07)	0.65	(0.06)	66	85	.001
1920	-0.11	(0.05)	0.14	(0.04)	11	5.3	.03
1924	-0.14	(0.05)	0.19	(0.04)	16	8.3	.006
1928	-0.06	(0.04)	0.10	(0.04)	5	2.1	.15
1932	0.022	(0.049)	0.023	(0.042)	0	.20	.66
1936	-0.005	(0.038)	0.039	(0.031)	0	.018	.90
1940	-0.20	(0.080)	0.26	(0.066)	12	6.4	.015
1944	-0.13	(0.024)	0.14	(0.021)	38	28.4	.001
1948	-0.39	(0.58)	0.46	(0.053)	25	15.7	.001
1952	-0.30	(0.04)	0.30	(0.04)	49	44.1	.001
1956	-0.18	(0.07)	0.21	(0.07)	11	5.7	.02
1960	-0.27	(0.04)	0.27	(0.04)	41	33.2	.001
1964	-0.24	(0.04)	0.24	(0.04)	43	30.3	.001
1968	-0.46	(0.04)	0.46	(0.04)	74	128	.001
1972	-0.29	(0.05)	0.29	(0.05)	41	32.1	.001
1976	-0.25	(0.04)	0.25	(0.04)	40	31.8	.001

MODEL: $\Delta H(T)_t = c_1 H(T)_t + c_2, \ \Delta H(T)_t > 0$

Year t	c_1	(error)	c_2	(error)	R^2 %	F	Sig.	N
1904	-0.11	(0.03)	0.13	(0.03)	34	11.9	.002	25
1908	-0.52	(0.08)	0.50	(0.06)	61	42.2	.001	31
1912	-0.21	(0.07)	0.23	(0.06)	33	8.3	.01	19
1916	-0.66	(0.04)	0.66	(0.03)	88	264	.001	36
1920	-0.12	(0.04)	0.13	(0.04)	32	8.3	.01	20
1924	-0.23	(0.07)	0.25	(0.06)	43	9.8	.008	15
1928	-0.15	(0.04)	0.15	(0.03)	47	15.3	.001	19
1932	0.0070	(0.037)	0.020	(0.030)	0	.03	.86	18
1936	0.00021	(0.056)	0.027	(0.045)	0	0	.9	18
1940	-0.34	(0.097)	0.38	(0.079)	25	12.2	.001	39
1944	-0.13	(0.03)	0.14	(0.03)	37	17.3	.001	31
1948	-0.55	(0.07)	0.57	(0.06)	85	64	.001	13
1952	-0.67	(0.13)	0.65	(0.11)	50	27	.001	38
1956	-0.29	(0.07)	0.29	(0.06)	67	18.2	.002	11
1960	-0.25	(0.05)	0.27	(0.04)	37	25	.001	44
1964	-0.17	(0.05)	0.18	(0.04)	23	11	.002	39
1968	-0.49	(0.05)	0.49	(0.05)	70	82	.001	37
1972	-0.29	(0.06)	0.27	(0.06)	42	19	.001	29
1976	-0.23	(0.04)	0.24	(0.04)	56	35	.001	29